Observations on Training a Neural Network with Pima Indians onset of diabetes dataset

Objective: Train a neural network model in PyTorch to predict if patient has diabetes or not

Dataset:

Input:

1. Number of times pregnant
2. Plasma glucose concentration at 2 hours in an oral glucose tolerance test
3. Diastolic blood pressure (mm Hg)
4. Triceps skin fold thickness (mm)
5. 2-hour serum insulin (µIU/ml)
6. Body mass index (weight in kg/(height in m)2)
7. Diabetes pedigree function
8. Age (years)

Target: Class label (0 or 1)

Training Procedure:

Initialized a basic feedforward neural network in PyTorch.
Used the Binary Cross Entropy (MSE) loss function since its a binary classification task.Used Adam optimizer with a learning rate of 1e-3.
Conducted training for 100 epochs with a batch size of 10.

Observations:

Convergence: The model demonstrated steady convergence across epochs, with loss declining gradually.
Performance: By the final epoch, the training loss was minimal, indicating a right fit to the data

Recommendations:

To improve generalization, introduce dropout layers or regularization techniques.
Consider normalizing the input data to aid the training process, especially if the range of numbers is large.

Conclusion: Using PyTorch, the neural network trained on the pima indian onset of diabetes dataset shows good performance with accuracy of 74% on unseen data during model evaluation on test dataset .The model's performance on unseen data suggests that it has learned meaningful patterns from the training data and generalizes well to new data.